

TECHNIQUES FOR IDENTIFYING MERGEABLE DATA

ABSTRACT OF THE DISCLOSURE

A system, method and article of manufacture for identifying mergeable data in a data processing system and, more particularly, for identifying correlated columns from one or more database tables. One embodiment comprises determining correlation attributes for a first column and a second column from one or more database tables. The correlation attributes describe for each column at least one of the column and content of the column. The correlation attributes from the first and second column are compared and similarities between the first and second column are identified on the basis of the comparison. Then, on the basis of the identified similarities, it is determined whether the first and second columns are correlated. Only if the columns are determined to be correlated, the first and second columns are merged.